JOURNAL OF BACTERIOLOGY, June 1989, p. 3564–3567
0021-9193/89/063564-04$02.00/0

Vol. 171, No. 6

A Protein Involved in Termination of Chromosome Replication in *Bacillus subtilis* Binds Specifically to the *terC* Site

P. J. LEWIS, M. T. SMITH, AND R. G. WAKE*

Department of Biochemistry, University of Sydney, Sydney, New South Wales 2006, Australia

Received 30 December 1988/Accepted 13 March 1989

The small basic protein encoded by the open reading frame adjacent to the *terC* site in the *Bacillus subtilis* chromosome and previously implicated in termination of the replication process was purified. Band retardation assays established that this protein (now called the replication terminator protein, encoded by the *rtp* gene) binds specifically to a 209-base-pair fragment of DNA within which *terC* is located.

Despite its importance, information on the molecular mechanism of termination of replication of the bacterial chromosome has lagged far behind that on the initiation and elongation phases of the replication cycle (7). There is evidence for the existence of sequence-specific terminators in the chromosomes of *Bacillus subtilis* 168 (11–13) and *Escherichia coli* (2, 4). In *B. subtilis*, the termination process appears to be less complex. A single termination site, called *terC*, is located approximately opposite the origin, *oriC*. The clockwise-moving replication fork reaches *terC* first and is arrested. The counterclockwise fork arrives a few (<5) minutes later, presumably to fuse with the arrested fork and to complete termination. Arrest of the clockwise fork at *terC* represents the first stage in the overall process of termination. The sequence of 1.3 kilobases of DNA spanning *terC* has been determined (1). Arrest occurs within this sequence in the vicinity of two imperfect inverted repeats (IRs; 47 and 48 nucleotides each, separated by 59 nucleotides) and just upstream of an open reading frame (ORF) encoding a small basic protein (M_r, 14,519). The ~200 base pairs (bp) of DNA spanning the IRs is referred to as the IR region (IRR). Experiments with strains in which portions of the sequence on either side of *terC* were deleted or modified have given information on the sequence requirements for clockwise fork arrest (10). In particular, disruption of the ORF by insertion of four extra nucleotides abolished fork arrest. It was proposed that arrest is dependent upon binding of the ORF-encoded protein to the IRR within which *terC* is located. In this report, we describe the purification of this protein from *B. subtilis* and the results of experiments which establish that the protein does bind specifically to the IRR as proposed. Thus, a protein required for termination of replication of a bacterial chromosome, now called the replication terminator (RT) protein, was identified.

Figure 1 summarizes the relevant features of the DNA sequence spanning *terC* in *B. subtilis*. The clockwise fork enters this sequence from the right and is arrested at *terC* just upstream of the ORF for the RT protein (designated the *rtp* gene) and in the vicinity of IRI and IRII, which make up the IRR (1). *Nde*I cuts this sequence between the ribosome-binding site and the initiation codon of the ORF; the *Hae*III site lies approximately 120 nucleotides downstream of the termination codon of the ORF. The 480-bp *Nde*I-*Hae*III portion was inserted, via a number of steps, into the multiple cloning site of *E. coli* expression vector pKK223-3 (Pharmacia, Uppsala, Sweden). In the resulting recombinant plasmid, pWS46, the ORF was positioned appropriately for controlled expression by the isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible *tac* promoter (Fig. 1, circular diagram). Figure 2A shows the result of sodium dodecyl sulfate-polyacrylamide gel electrophoresis (8) of the total protein in uninduced (lane 1) and induced (lane 2) cultures of *E. coli* XL1-blue (Stratagene) harboring pWS46 (see below for the induction procedure). The latter shows an additional and prominent species with an approximate M_r of 14,500. It constitutes about 5% of the total protein and is consistent with its reflecting overexpression of the RT protein. (The same 14,500-M_r species was by far the major protein produced following use of an in vitro expression system with pWS46 DNA [data not shown].) Advantage was taken of the high calculated pI (9.2) of the RT protein to purify the overexpressed protein by cation-exchange chromatography. The overall procedure was as follows. After growth to the mid-exponential phase (A_{600}, 0.7) in 2× TY broth–ampicillin (100 μg/ml) at 37°C, IPTG was added to 0.5 mM and incubation was continued for 1 h before harvesting. Cells from a 500-ml culture were suspended in 10 ml of buffer A (25 mM Tris hydrochloride [pH 7.8], 250 mM sucrose, 1 mM phenylmethylsulfonyl fluoride, 2 mM EDTA), recentrifuged, frozen in liquid N_2, and stored at −80°C. Cells were disrupted with a French press at −27°C and suspended in 5 ml of buffer B (50 mM Tris hydrochloride [pH 7.8], 2 mM EDTA, 0.5 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride) at 0°C. This was centrifuged at 10,000 × *g* and −2°C for 30 min, and the thick, clear amber supernatant was kept. Nucleic acids were removed by stepwise (100 μl) addition of protamine sulfate (10 mg/ml in 50 mM Tris hydrochloride [pH 7.8]) at 4°C to a final level of 20 mg. The suspension was centrifuged (3,000 × *g*) at 4°C for 5 min, and the clear supernatant was analyzed by agarose gel electrophoresis and ethidium bromide staining. If nucleic acids were still obvious, a further 10 mg of protamine sulfate was added in the same way and the suspension was recentrifuged. At this stage, the extract was clear and colorless and much less viscous than the original extract. It was dialyzed twice against 1 liter of 2× buffer C [50 mM 2-(*N*-morpholino) ethanesulfonic acid buffer (pH 6.0), 0.5 mM dithiothreitol] at 4°C overnight. Cation-exchange chromatography was performed with a fast-protein liquid chromatography system (Pharmacia) with a Mono S column (HR 5/5). The total extract was loaded onto the column at room temperature and eluted with a gradient of NaCl (salt) in buffer C. The RT

* Corresponding author.

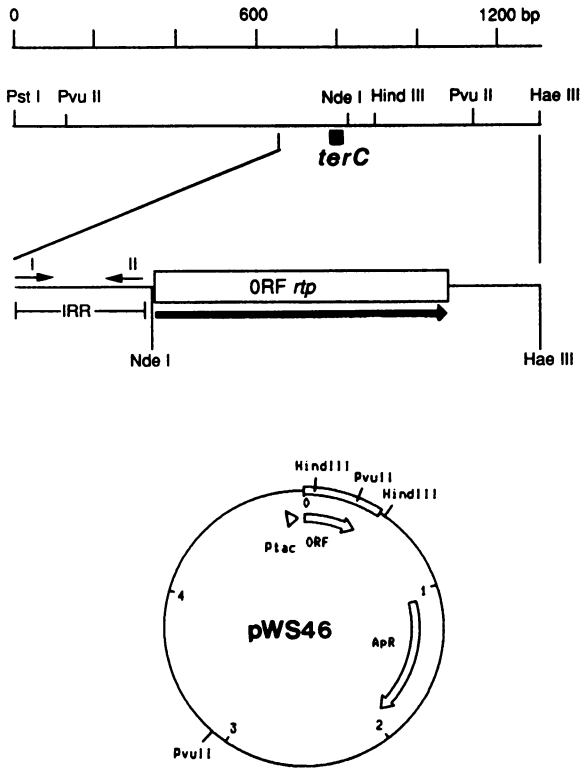

FIG. 1. Sequence features of 1.3 kilobases of DNA spanning *terC* in *B. subtilis*. The clockwise replication fork enters the segment of the chromosome shown as a restriction map from the right and stops at *terC* (■). Relevant sequence features of an enlarged portion of the map are shown below it. These comprise two IRs (I and II, arrows) just upstream of an ORF which encodes the RT protein (*rtp* gene) and whose direction of transcription is shown by the thick arrow. The 480-bp *Nde*I-*Hae*III fragment spanning the ORF was inserted into the multiple cloning site of *E. coli* expression vector pKK223-3 to give plasmid pWS46 (lower section), in which the ORF was expressed under control of the *tac* promoter (Ptac). The cloned insert is shown as an open box; the vector-encoded ampicillin resistance gene (ApR) is also shown.

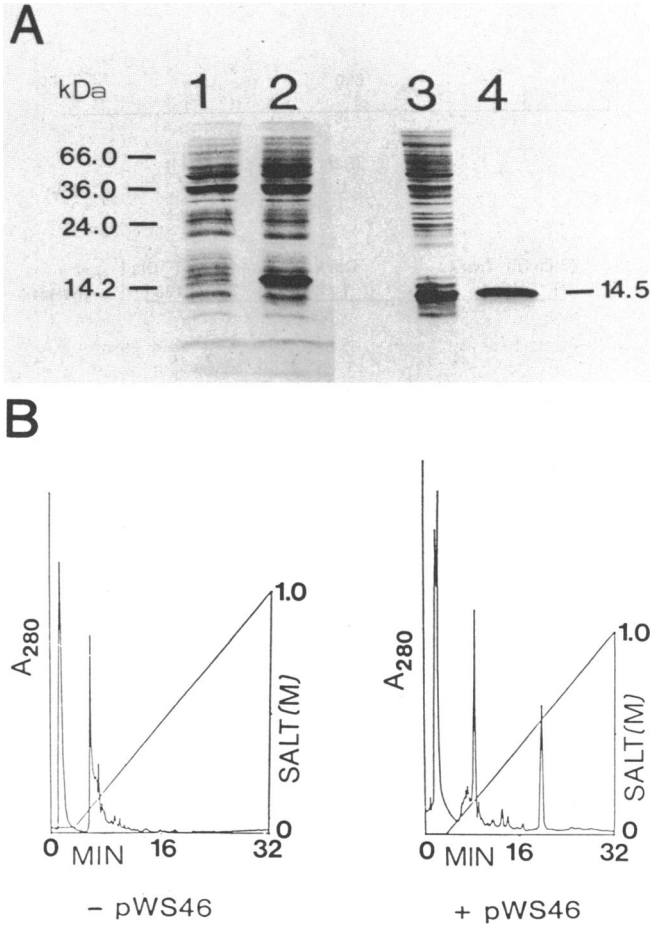

FIG. 2. Overexpression and purification of the RT protein. Panel A (lanes 1 and 2) shows the result of sodium dodecyl sulfate-polyacrylamide gel electrophoresis (15% gel; reference 8) of the total cell protein (washed cells dissolved directly in loading buffer containing 0.125 M Tris hydrochloride [pH 6.8], 4% sodium dodecyl sulfate, 20% glycerol, and 10% 2-mercaptoethanol) from control (uninduced, no IPTG) and induced cultures, respectively. Lanes 3 and 4 compare the total protein from an induced culture and the purified RT protein, respectively. The molecular mass markers in panel A apply only to lanes 1 and 2. kDa, Kilodaltons. Panel B compares the protein elutions (A_{280}) from the fast-protein liquid chromatography column for extracts from a control culture of *E. coli* XL1-blue (left) and a culture of the same strain harboring pWS46 (right), both treated with IPTG in the same way. The control culture was grown without ampicillin.

protein eluted as a sharp peak at 0.56 M NaCl. The material in the peak was rechromatographed, concentrated in a Centricon 10 microconcentrator (Amicon Corp., Lexington, Mass.), and stored in buffer C at 4°C. (Its DNA binding was stable for at least 5 weeks under these conditions [see below].) The yield was ~150 μg of purified protein per 500-ml induced culture. Figure 2B compares the protein elution profiles (from the cation-exchange fast-protein liquid chromatography column) of nucleic acid-free extracts prepared from IPTG-induced cultures of *E. coli* with or without pWS46. The former (+pWS46) showed a prominent, sharp peak eluting at 0.56 M salt, comprising the RT protein. It was completely absent from the control (−pWS46) extract. The final purified product comprised ~90% of the 14,500-M_r species (Fig. 2A, lane 4).

Band retardation assays (3) were performed to explore the possibility of binding of the RT protein to the IRR as proposed. Digestion of plasmid pPL-1 with appropriate enzymes yielded a small fragment of 219 bp containing the IRR and another of 134 bp from the adjoining region to the right (Fig. 3A). The vector was present in a fragment of 3.0 kilobases. To examine the effect of the purified RT protein on the electrophoretic migration of these fragments, it was diluted in 50 mM Tris hydrochloride (pH 7.8)–50% glycerol to the appropriate concentration and mixed with digested DNA in TGMK buffer (50 mM Tris hydrochloride [pH 7.8], 50% glycerol, 10 mM $MgCl_2$, 50 mM KCl) in a total volume of 6.0 μl. Care was taken to ensure only gentle mixing of the components with the pipette tip after addition of the protein to the DNA. (In the present experiments, each sample contained 0.14 pmol of each individual DNA fragment and 0 to 11.4 pmol of the RT protein monomer.) The mixtures were incubated at 37°C for 10 min and quenched on ice. Bromophenol blue (0.07%) was added before loading of the complete sample onto a 4% wide-range agarose gel (Sigma Chemical Co., St. Louis, Mo.) in electrophoresis buffer (36 mM Tris, 30 mM NaH_2PO_4, 10 mM EDTA [pH 7.5]). Electrophoresis was performed in a horizontal submerged

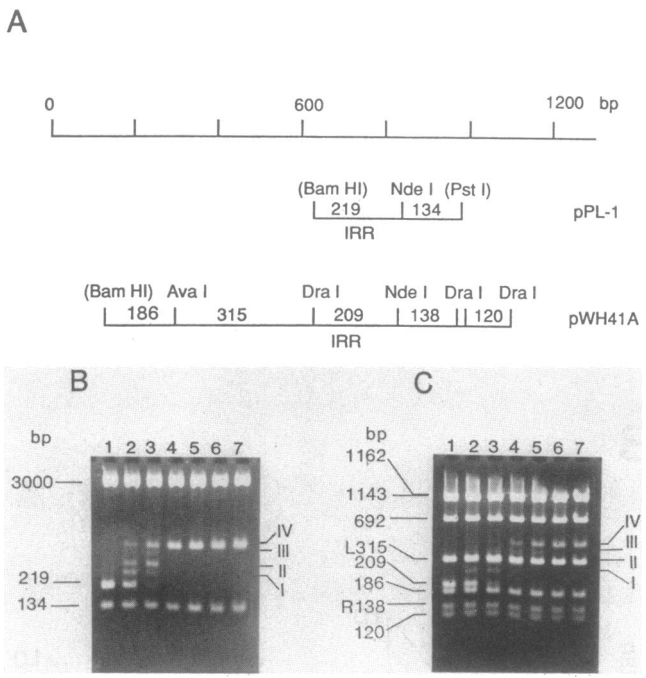

FIG. 3. Specific binding of the RT protein to the DNA fragment that contains the IRR. The scale in panel A refers to the map at the top of Fig. 1. pPL-1 and pWH41A are plasmids containing *B. subtilis* inserts as shown. The brackets enclosing three of the end restriction sites indicate that they lie in the vector portion of the plasmid and just outside the *B. subtilis* insert. The 134-, 186-, and 219-bp segments include <8 bp of vector DNA. Each of these plasmids was used to provide a collection of small fragments from the chromosomal region in the vicinity of the IRR for use in band retardation assays with the RT protein. Panel B shows the effects of increasing levels of RT protein on the migration of fragments derived from pPL-1. The molar ratios (*r*) of RT protein (monomer of 14.5 kilodaltons) in individual DNA fragments were 0, 7, 15, 20, 27, 54, and 80 for lanes 1 to 7, respectively. The 3,000-bp fragment represents the vector. Panel C shows the effects of increasing levels of RT protein on the migration of fragments derived from pWH41A; r = 0, 3, 7, 14, 20, 27, and 54 for lanes 1 to 7, respectively. The 692-, 1,143-, and 1,162-bp fragments are vector DNA. The L315 fragment lies immediately to the left of the IRR fragment in the chromosome, and the R138 fragment lies to its right.

mini-gel apparatus (gel dimensions, 8.0 by 5.4 by 0.5 cm) at 62 mA for 1.5 h in a cold room; the gels were stained with ethidium bromide before photography.

Figure 3B (lane 1) shows the electrophoretic resolution of the three species derived from pPL-1; lanes 2 to 7 show the effects of increasing amounts of the purified RT protein on their migration. The 219-bp fragment was converted to a series of more slowly moving discrete bands (I to IV) at the lower protein-DNA molar ratios (r = 7 to 15; Fig. 3, legend) and just the slowest species (IV) at the highest ratio (lane 7; r = 80). The 134-bp fragment remained unaltered throughout. Binding of the RT protein to the 219-bp fragment to produce a ladder of retarded bands is somewhat similar to the situation with the *lac* repressor-operator interaction at high protein-DNA ratios (3). The latter was explained by an increase in the number of *lac* repressor tetramers bound per DNA fragment as the ratio increased.

The specificity of the binding of the RT protein to the IRR was examined more stringently in the experiment shown in Fig. 3C. Digestion of plasmid pWH41A with appropriate enzymes yielded five *terC* region fragments (120 to 315 bp) (Fig. 3A). In addition to the 209-bp IRR fragment, there were two (120 and 138 bp) from one side of the IRR and two (186 and 315 bp) from the other side. Most of the vector was distributed among three larger fragments of 692, 1,143, and 1,162 bp. All of these fragments, except the two largest ones, resolved from one another under the electrophoretic conditions used (Fig. 3C, lane 1). Addition of increasing amounts of the RT protein (lanes 2 to 7; r = 3 to 54) caused selective retardation of the 209-bp IRR fragment. There was no evidence for nonspecific interaction of the protein with any fragment when the protein-DNA ratio was twice that needed to completely convert the 209-bp fragment to the slowest derived species (IV) of the ladder (compare lanes 6 and 7). Lane 4 shows that all of the 209-bp fragment was complexed to protein when $r \simeq 14$. To ascertain whether the RT protein bound to each of the IRs within the IRR, synthetic oligonucleotides equivalent to IRI and IRII, flanked by *Bam*HI and *Eco*RI sites, were cloned in the Bluescript vector KSM13− (Stratagene) to give pIRI and pIRII, respectively. Band retardation assays using *Bam*HI-*Eco*RI-cut pIRI and pIRII showed that the RT protein bound to IRI and IRII separately and with approximately identical affinities (data not shown). Furthermore, in each case at low protein-DNA ratios only two retarded species per IR were observed. This raises the possibility that the four retarded bands (I to IV) seen in Fig. 3B arose through the binding of two sequential units of the RT protein to each IRI and IRII.

The data presented here firmly establish that the RT protein of *B. subtilis* encoded by the ORF adjacent to *terC* binds specifically to the segment of DNA that encompasses the IRR and in which *terC* is located. This finding is strong support for the proposal that binding of the RT protein within the IRR is directly involved in arresting the clockwise fork in this region at the *terC* site.

The original basis for suggesting that the RT protein binds to the IRR was threefold. (i) The presence of the long imperfect IRs, IRI and IRII, in the vicinity of *terC* was a clearly recognizable feature of the region (1), (ii) deletion of IRI abolished fork arrest (10), and (iii) deletion or disruption of the ORF for the RT protein also abolished fork arrest (10). It remains to be established whether or not binding of the RT protein to the IRR per se is sufficient to cause arrest. When the IRs close to the *terC* site in *B. subtilis* were identified, the terminus of *E. coli* plasmid R6K was also examined and found to contain two 20-nucleotide imperfect IRs separated by 73 nucleotides (1). These R6K repeats (called *terR1* and *terR2*) have since been studied in more detail, and each has been shown to function separately as a block to replication fork movement in one or the other direction (6). There is evidence to suggest that *terR1* and *terR2* of R6K show significant sequence homology to the *E. coli* chromosomal terminators. It is possible that the recently identified *trans*-acting factor (defined by the *tus* gene) needed for termination of replication in *E. coli* (5) is a protein, analogous in function to the RT protein identified here, that binds separately to each of the *E. coli* chromosomal terminators, as well as to *terR1* and *terR2* of R6K.

Knowledge of the mode of binding of the RT protein to the IRR in *B. subtilis* will be of some significance. Of the many DNA-binding proteins identified, it has a unique function. Also, it is likely that its precise folded structure must be stringently maintained for it to function in termination of replication. Examination of the ORF for the RT protein in another strain of *B. subtilis*, W23 (all of the data presented here relate to strain 168), showed 22 nucleotide changes in the ORF but none in the encoded amino acid sequence (9).

This work received financial support from the Australian Research Council and the Sir Zelman Cowen Universities Fund.

LITERATURE CITED

1. **Carrigan, C. M., J. A. Haarsma, M. T. Smith, and R. G. Wake.** 1987. Sequence features of the replication terminus of the *Bacillus subtilis* chromosome. Nucleic Acids Res. **15**:8501–8509.
2. **de Massy, B., S. Bejar, J. Louarn, J.-M. Louarn, and J.-P. Bouche.** 1987. Inhibition of replication forks exiting the terminus region of the *Escherichia coli* chromosome occurs at two loci separated by 5 min. Proc. Natl. Acad. Sci. USA **84**:1759–1763.
3. **Freid, M., and D. Crothers.** 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9**:6505–6525.
4. **Hill, T. M., J. M. Henson, and P. L. Kuempel.** 1987. The terminus region of the *Escherichia coli* chromosome contains two separate loci that exhibit polar inhibition of replication. Proc. Natl. Acad. Sci. USA **84**:1754–1758.
5. **Hill, T. M., B. J. Kopp, and P. L. Kuempel.** 1988. Termination of DNA replication in *Escherichia coli* requires a *trans*-acting factor. J. Bacteriol. **170**:662–668.
6. **Horiuchi, T., and M. Hidaka.** 1988. Core sequence of two separable terminus sites of the R6K plasmid that exhibit polar inhibition of replication is a 20 bp inverted repeat. Cell **54**: 515–523.
7. **Kornberg, A.** 1988. DNA replication. J. Biol. Chem. **263**:1–4.
8. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
9. **Lewis, P. J., and R. G. Wake.** 1989. DNA and protein sequence conservation at the replication terminus in *Bacillus subtilis* 168 and W23. J. Bacteriol. **171**:1402–1408.
10. **Smith, M. T., and R. G. Wake.** 1988. DNA sequence requirements for replication fork arrest at *terC* in *Bacillus subtilis*. J. Bacteriol. **170**:4083–4090.
11. **Weiss, A. S., I. K. Hariharan, and R. G. Wake.** 1981. Analysis of the terminus region of the *Bacillus subtilis* chromosome. Nature (London) **293**:673–675.
12. **Weiss, A. S., and R. G. Wake.** 1983. Restriction map of DNA spanning the replication terminus of the *Bacillus subtilis* chromosome. J. Mol. Biol. **171**:119–137.
13. **Weiss, A. S., and R. G. Wake.** 1984. A unique DNA intermediate associated with termination of chromosome replication in *Bacillus subtilis*. Cell **39**:683–689.